PCT/NZ03/00119

Intellectual Property Office of New Zealand
Te Pou Rāhui Hanga Hou

REC'D 15 JUL 2003
WIPO PCT

CERTIFICATE

This certificate is issued in support of an application for Patent registration in a country outside New Zealand pursuant to the Patents Act 1953 and the Regulations thereunder.

I hereby certify that annexed is a true copy of the Provisional Specification as filed on 10 June 2002 with an application for Letters Patent number 519469 made by AGRESEARCH LIMITED.

Dated 3 July 2003.

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

Neville Harris

Neville Harris
Commissioner of Patents

COMMISSIONER OF PATENTS, TRADE MARKS & DESIGNS • NEW ZEALAND •

Best Available Copy

519469

PATENTS FORM NO. 4

Appln Fee: $50.00

James & Wells ref: 31161/39

PATENTS ACT 1953

PROVISIONAL SPECIFICATION

IMPROVEMENTS IN AND RELATING TO A METHOD OF DNA TESTING

I/WE AGRESEARCH LIMITED a New Zealand Company of East Street, Ruakura Campus, Hamilton, New Zealand

do hereby declare this invention to be described in the following statement:

IMPROVEMENTS IN AND RELATING TO A METHOD OF DNA TESTING

TECHNICAL FIELD

This invention relations to improvements in and relating to a method of DNA testing. In particular this invention relates to nucleic acid sequences of *Mycobacterium paratuberculosis* and their use in a method for identifying different strains of *M. paratuberculosis* and distinguishing strains of *M. paratuberculosis* from other mycobacterial species. This invention also provides an aid in the diagnosis of diseases caused by Mycobacterial species in human and animal medical practice.

BACKGROUND ART

Mycobacteria are rod-shaped, acid-fast, aerobic bacilli that do not form spores. A moderate number of slow-growing mycobacterial species are major pathogens for humans and/or animals. For example, paratuberculosis is a very widespread animal health problem which causes major economic losses in farming of ruminant animals particularly in the dairy industry. The development of robust diagnostic tests to distinguish different mycobacterial species and to characterise subspecies, groups and types of related strains within a species is of prime importance.

Paratuberculosis or Johne's disease is a chronic granulomatous enteritis that can affect all domestic and wild ruminants causing reduced food intake, weight loss and death. The disease is present in most countries and results in significant production losses. The causative organism, *Mycobacterium avium* subsp. *paratuberculosis* (basonym *M. paratuberculosis*) (Harris *et al.*, 2001) has also been implicated as the etiologic agent of Crohn's disease in humans and is a member of the MAI complex, a group of closely related species which includes *Mycobacterium intracellulare* and all subspecies of *M. avium*. For taxonomic purposes, *M. avium* is divided into the three subspecies *M. avium* subsp. *avium* (although in most publications this subspecies is still referred to as *M. avium*), *M. avium* subsp. *paratuberculosis* and *M. avium* subsp. *silvaticum* (Thorel

et al., 1990). While *M. paratuberculosis* appears to be an obligate pathogen, closely-related organisms of the MAI complex that share many common antigens with *M. paratuberculosis* are widespread throughout the environment. Exposure of animals to these environmental organisms is probably responsible for the lack of sensitivity and specificity of antigen-based diagnostic tests for *M. paratuberculosis*. Other problems that have made this disease particularly difficult to control are the very slow growth of the organism on artificial culture, and the ability of the organism to survive in many animals for years without causing any overt disease (Chiodini et al., 1984; Harris and Barletta, 2001).

Two recent discoveries have shown that the spread of *M. paratuberculosis* may be more complicated than previously believed and emphasise the need for the development of new diagnostic tools. First, the organism has been reported to survive normal milk pasteurisation (Grant *et al.*, 2002). Since pasteurised milk is widely consumed in many countries, this survival provides a route by which large sections of the population can be exposed to this obligate pathogen and supports the case of those who claim that it causes Crohn's disease in humans (Hermon-Taylor *et al.*, 2000; Harris and Barletta, 2001). Second, *M. paratuberculosis* has also been isolated in the United Kingdom from common wild non-ruminant animals such as rabbits, foxes, stoats and crows (Beard *et al.*, 2001). This finding complicates epidemiological studies, as previously it had been believed that spread of the disease occurred only from ruminants, either directly from one animal to another or through infected milk or by grazing on pasture infected by organisms shed from another infected ruminant (Chiodini *et al.*, 1984).

The first significant molecular biological development in the study of *M. paratuberculosis* was the discovery of multiple copies of an insertion sequence IS*900* (Collins *et al.*, 1989; Green *et al.*, 1989). This sequence has been found to be specific for *M. paratuberculosis* and is now widely used as the basis for diagnostic tests that

use DNA amplification (Collins *et al.*, 1993; Fang *et al.*, 2002). Related insertion sequences have been found in other members of the MAI complex (Kunze *et al.*, 1991; Englund et al., 2002) and the finding that the most recently discovered sequence is 94% identical to IS*900* has raised doubts about the specificity of tests based on parts of the IS*900* sequence (Englund *et al.*, 2002). There would be advantages in having a range of sequences that have a reasonable probability of being specific to *M. paratuberculosis* so that new tests could be widely trialled to determine which sequences are truly unique to this species. Sequencing of the genomes of both an *M. paratuberculosis* and an *M. avium* subsp. *avium* strain is currently in progress and a range of sequences that might differ between these two strains have been identified (Bannantine, 2002). Whether all these differences are real cannot be determined until the sequencing of both genomes is completed but even then the genetic diversity of different MAI strains is such (Falkinham, 1999) that it will be some years before the degree of specificity of these sequences can be determined for a wide range of strains in the different subspecies.

Isolates of *M. paratuberculosis* were first characterised into cattle and sheep types in 1990 (Collins *et al.*, 1990) on the basis of restriction fragment length polymorphisms (RFLPs) of the insertion sequence IS*900* and this largely correlates with the difficulty of primary isolation of sheep types (Collins *et al.*, 1990, Pavlik *et al.*, 1999). The distinction into cattle and sheep types is epidemiologically useful, as cattle and sheep are preferentially infected with their named types while other ruminant species such as deer and goats appear to be infected more easily with either type (Collins *et al.*, 1990; de Lisle *et al.*, 1993; Pavlik *et al.*, 1999; Whittington *et al.*, 2000). Sheep strains from Canada (Collins *et al.*, 1990) and subsequently from South Africa (de Lisle *et al.*, 1992) and Iceland (de Lisle *et al.*, 1993) were found to have RFLP patterns that clustered in a group that was different from that of cattle types and other sheep types and were classified as belonging to a third or intermediate type. A careful comparison of members of these three RFLP types revealed that the pattern of the intermediate type

was more closely related to patterns of the other sheep type than to patterns of the cattle type (Pavlik *et al.*, 1999) and for this reason, and also because of its epidemiological association with sheep, this intermediate type is better referred to as a variant or second sheep type.

At present, DNA amplification testing for paratuberculosis where both cattle and sheep types are potentially present, involves a PCR assay based on IS*900* to confirm the presence of *M. paratuberculosis* followed by a PCR based on IS*1311* whose product is then subjected to restriction endonuclease analysis (Whittington *et al.*, 2000). This two-step PCR analysis approach is performed because IS*1311* is not unique for *M. paratuberculosis* and is also found in *M. avium* subsp. *avium* (Collins *et al.*, 1997), but some copies of IS*1311* in *M. paratuberculosis* have polymorphisms that are specific for the cattle and sheep types and the polymorphisms can be detected by digesting the IS*1311* PCR product with appropriate restriction enzymes (Marsh *et al.*, 1999).

All references, including any patents or patent applications cited in this specification are hereby incorporated by reference. No admission is made that any reference constitutes prior art. The discussion of the references states what their authors assert, and the applicants reserve the right to challenge the accuracy and pertinency of the cited documents. It will be clearly understood that, although a number of prior art publications are referred to herein, this reference does not constitute an admission that any of these documents form part of the common general knowledge in the art, in New Zealand or in any other country.

It is acknowledged that the term 'comprise' may, under varying jurisdictions, be attributed with either an exclusive or an inclusive meaning. For the purpose of this specification, and unless otherwise noted, the term 'comprise' shall have an inclusive meaning - i.e. that it will be taken to mean an inclusion of not only the listed components it directly references, but also other non-specified components or elements. This rationale will also be used when the term 'comprised' or 'comprising' is

used in relation to one or more steps in a method or process.

It is an object of the present invention to address the foregoing problems or at least to provide the public with a useful choice.

Further aspects and advantages of the present invention will become apparent from the ensuing description which is given by way of example only.

SUMMARY OF INVENTION

The present invention relates to the discovery of a DNA sequence in sheep types of *M. paratuberculosis* that differs from the homologous sequence in cattle types of *M. paratuberculosis.* The invention also provides a nucleic acid amplification technique based on these differences that can be used to distinguish strains of the cattle type from strains of both the sheep types of *M. paratuberculosis*. The invention also relates to use of these sequences in a nucleic acid amplification technique to distinguish all strains of *M. paratuberculosis* from other strains of the MAI complex and from strains of the *M. tuberculosis* complex.

DISCLOSURE OF INVENTION

According to a first aspect of the present invention there is provided a nucleic acid molecule of a sheep type of *M. paratuberculosis* said molecule comprising SEQ ID NO. 1 or its complement.

According to a second aspect of the present invention there is provided a probe comprising SEQ ID NO.1 or its complement.

According to a third aspect of the present invention there is provided a probe comprising at least 6 contiguous nucleotides selected from nucleotides 1 – 35 of SEQ ID NO. 1 or its complement.

Preferably, the probe substantially as described above may include at least 10-12 contiguous nucleotides selected from nucleotides 1 – 35 of SEQ ID NO. 1 or its complement.

More preferably the probe substantially as described above may include more than 20 contiguous nucleotides selected from nucleotides 1 – 35 of SEQ ID NO. 1 or its complement.

According to a fourth aspect of the present invention there is provided a probe including at least 6 contiguous nucleotides selected from nucleotides 230 – 260 of SEQ ID NO. 1 or its complement.

Preferably the probe substantially as described above may include 10-12 contiguous nucleotides selected from nucleotides 230 – 260 of SEQ ID NO. 1 or its complement.

More preferably the probe substantially as described above may include more than 20 contiguous nucleotides selected from nucleotides 230 – 260 of SEQ ID NO. 1 or its complement.

According to a fifth aspect of the present invention there is provided a use of SEQ ID NO 1 or its complement for detecting the presence of sheep types of *M. paratuberculosis*.

According to a sixth aspect of the present invention there is provided a use of SEQ ID NO 2 or its complement for detecting the presence of cattle types of *M. paratuberculosis*.

According to a seventh aspect of the present invention there is provided a method of distinguishing between cattle and sheep types of *M. paratuberculosis* by comparing differences in nucleotide sequences of SEQ ID NO. 1 and SEQ ID NO. 2 or their complements.

According to a eighth aspect of the present invention there is provided a method of detecting the presence of *M. paratuberculosis* in a sample via a nucleic acid amplification technique said method comprising the steps of:

a) taking a sample from an animal or any other source

b) extracting nucleic acids from the sample or culturing mycobacteria from the sample and extracting nucleic acids from the mycobacterial culture

c) performing a nucleic acid amplification technique.

d) determining the identity of the amplification product.

Preferably animals may include cattle, sheep, deer, goats, ferrets, rabbits and humans but should not be limited thereto.

According to a ninth aspect of the present invention there is provided a method substantially as described above wherein step c) and/or d) of the method comprise identifying the presence of at least 6 nucleotides of the nucleic acid molecule comprising SEQ ID NO. 1 or its complement.

According to a tenth aspect of the present invention there is provided a method substantially as described above wherein step c) and/or d) of the method comprise identifying the presence of 10-12 nucleotides of the nucleic acid molecule comprising SEQ ID NO. 1 or its complement.

According to an eleventh aspect of the present invention there is provided a method substantially as described above wherein step c) and/or d) of the method comprise identifying the presence of at least 15 nucleotides of the nucleic acid molecule comprising SEQ ID NO. 1 or its complement.

According to a twelfth aspect of the present invention there is provided a method substantially as described above wherein step c) and/or d) of the method comprise

identifying the presence of approximately 20 nucleotides of the nucleic acid molecule comprising SEQ ID NO. 1 or its complement.

According to a thirteenth aspect of the present invention there is provided a method substantially as described above wherein step c) utilizes one oligonucleotide primer complementary to at least 6 contiguous nucleotides of SEQ ID NO. 1 or its complement and one oligonucleotide primer complementary to at least 6 nucleotides of IS*900* or its complement.

According to a fourteenth aspect of the present invention there is provided a method substantially as described above wherein step c) utilizes one oligonucleotide primer complementary to 10-12 contiguous nucleotides of SEQ ID NO. 1 or its complement and one oligonucleotide primer complementary to 10-12 nucleotides of IS*900* or its complement.

According to a fifteenth aspect of the present invention there is provided a method substantially as described above wherein step c) utilizes one oligonucleotide primer complementary to approximately 15 contiguous nucleotides of SEQ ID NO. 1 or its complement and one oligonucleotide primer complementary to approximately 15 nucleotides of IS*900* or its complement.

According to a sixteenth aspect of the present invention there is provided a method substantially as described above wherein step c) utilizes one oligonucleotide primer complementary to approximately 20 contiguous nucleotides of SEQ ID NO. 1 or its complement and one oligonucleotide primer complementary to approximately 20 nucleotides of IS*900* or its complement.

According to a seventeenth aspect of the present invention there is provided a method substantially as described above wherein step c) and/or d) of the method comprise identifying the presence of at least 6 nucleotides of the nucleic acid molecule comprising SEQ ID NO. 2 or its complement.

According to an eighteenth aspect of the present invention there is provided a method substantially as described above wherein step c) and/or d) of the method comprise identifying the presence of 10-12 nucleotides of the nucleic acid molecule comprising SEQ ID NO. 2 or its complement.

According to a nineteenth aspect of the present invention there is provided a method substantially as described above wherein step c) and/or d) of the method comprise identifying the presence of at least 15 nucleotides of the nucleic acid molecule comprising SEQ ID NO. 2 or its complement.

According to a twentieth aspect of the present invention there is provided a method substantially as described above wherein step c) and/or d) of the method comprise identifying the presence of approximately 20 nucleotides of the nucleic acid molecule comprising SEQ ID NO. 2 or its complement.

According to a twenty-first aspect of the present invention there is provided a method substantially as described above wherein step c) utilizes one oligonucleotide primer complementary to at least 6 contiguous nucleotides of SEQ ID NO. 2 or its complement and one oligonucleotide primer complementary to at least 6 nucleotides of IS*900* or its complement.

According to a twenty-second aspect of the present invention there is provided a method substantially as described above wherein step c) utilizes one oligonucleotide primer complementary to 10-12 contiguous nucleotides of SEQ ID NO. 2 or its complement and one oligonucleotide primer complementary to 10-12 nucleotides of IS*900* or its complement.

According to a twenty-third aspect of the present invention there is provided a method substantially as described above wherein step c) utilizes one oligonucleotide primer complementary to approximately 15 contiguous nucleotides of SEQ ID NO. 2 or its

complement and one oligonucleotide primer complementary to approximately 15 nucleotides of IS*900* or its complement.

According to a twenty-fourth aspect of the present invention there is provided a method substantially as described above wherein step c) utilizes one oligonucleotide primer complementary to approximately 20 contiguous nucleotides of SEQ ID NO. 2 or its complement and one oligonucleotide primer complementary to approximately 20 nucleotides of IS*900* or its complement.

According to a twenty-fifth aspect of the present invention there is provided use of SEQ ID NO.1 and SEQ ID NO. 2 substantially as described above to determine whether a strain of either a sheep type or a cattle type of *M. paratuberculosis* is present in a sample.

According to a twenty-sixth aspect of the present invention there is provided a use of SEQ ID NO.1 to distinguish any strain of *M. paratuberculosis* from any other strain of the MAI complex which may be present in a sample.

According to a twenty-seventh aspect of the present invention there is provided a use of SEQ ID NO.2 to distinguish any strain of *M. paratuberculosis* from any other strain of the MAI complex which may be present in a sample.

According to a twenty-eighth aspect of the present invention there is provided a use of SEQ ID NO.1 to distinguish any strain of *M. paratuberculosis* from any strain of the *M. tuberculosis* complex which may be present in a sample.

According to a twenty-ninth aspect of the present invention there is provided a use of SEQ ID NO.2 to distinguish any strain of *M. paratuberculosis* from any strain of the *M. tuberculosis* complex which may be present in a sample.

According to a thirtieth aspect of the present invention there is provided a use of SEQ ID NO. 1 to detect the presence of *M. paratuberculosis* as a causative agent of Johne's disease or Crohn's disease.

According to a thirty-first aspect of the present invention there is provided a use of SEQ ID NO. 2 to detect the presence of *M. paratuberculosis* as a causative agent of Johne's disease or Crohn's disease.

It will be appreciated by those skilled in the art such that knowledge and use of the nucleotide sequences of SEQ ID NO. 1 and SEQ ID NO. 2 will be useful as an aid in the diagnosis of Johne's and Crohn's diseases.

The term "sheep type of *M. paratuberculosis*" as used herein refers to a strain of *M. paratuberculosis* which preferentially infects sheep but also may infect other species for example deer, goats and humans but does not preferentially infect cattle.

The term "cattle type of *M. paratuberculosis*" as used herein refers to a strain of *M. paratuberculosis* which preferentially infects cattle but also may infect other species for example deer, goats and humans but does not preferentially infect sheep.

The term "IS900" as used herein refers to a known DNA sequence containing genes unique to *M. paratuberculosis* which may be used to detect *M. paratuberculosis* species.

The term "MAI complex" as used herein refers to a group of closely related species which includes *Mycobacterium intracellulare* and all subspecies of *M. avium*.

The term "*M. tuberculosis* complex" as used herein refers to the group of organisms which cause tuberculosis in mammals comprise the following species: *M. tuberculosis, M. bovis, Mycobacterium bovis* subsp. *caprae, Mycobacterium microti* and *Mycobacterium canettii*.

"Probes" are single-stranded nucleic acid molecules with a known nucleotide sequence which is labelled in some way (for example, radioactively, fluorescently or immunologically), which are used to find and mark a target DNA or RNA sequence by hybridizing to it.

"Primers" are short nucleic acids, preferably DNA oligonucleotides 15 nucleotides or more in length, which are annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand; they can then be extended along the target DNA strand by a polymerase, preferably a DNA polymerase. Primer pairs can be used for amplification of a nucleic acid sequence, e.g. by the polymerase chain reaction (PCR) or other nucleic acid amplification methods well known in the art. PCR-primer pairs can be derived from the sequence of a nucleic acid according to the present invention, for example, by using computer programs intended for that purpose such as Primer (Version 0.5© 1991, Whitehead Institute for Biomedical Research, Cambridge, MA).

The term "complement," as used herein is best illustrated by the following example. For the sequence 5' AGGACC 3', the complement would be 3' TCCTGG 5'.

The term "nucleic acid molecule" as used herein may be an RNA, cRNA, genomic DNA or cDNA molecule, and may be single- or doublestranded. The nucleic acid molecule may also optionally comprise one or more synthetic, non-natural or altered nucleotide bases, or combinations thereof.

The term "nucleic acid amplification technique" as used herein may generally be considered to refer to polymerase chain reaction or PCR however, it may equally refer to other equivalent techniques for amplifying nucleic acids known to those skilled in the art.

The term "polymerase chain reaction or PCR" as used herein refers to a system for *in vitro* amplification of DNA. Two synthetic oligonucleotide primers, which are

complementary to two regions of the target DNA (one for each strand) to be amplified, are added to the target DNA (that need not be pure), in the presence of excess deoxynucleotides and Taq polymerase, a heat-stable DNA polymerase. In a series of temperature cycles, the target DNA is repeatedly denatured, annealed to the primers (typically at 50-60°C) and a daughter strand extended from the primers. As the daughter strands themselves act as templates for subsequent cycles, DNA fragments matching both primers are amplified exponentially.

The detection of the amplified nucleic acid may be by any of a wide range of techniques known to those skilled in the art, including but not limited to size separation techniques such as gel electrophoresis, probe detection systems either on solid supports or in solution and DNA microarray techniques.

BRIEF DESCRIPTION OF DRAWINGS

Further aspects of the present invention will become apparent from the following description which is given by way of example only and with reference to the accompanying drawings in which:

Figure 1 Shows PCR products from DNA of two sheep types and two cattle types of *M. paratuberculosis* performed with oligonucleotide primers DMC136 and DMC137 at an annealing temperature of 50°C;

Figure 2 Shows Alignment of homologous DNA sequences from cattle and sheep types of *M. paratuberculosis*. Identical nucleotides in both sequences are shaded, arrows indicate the identity and direction of the oligonucleotide primers used, the tandem DNA sequence present in the sheep type is shown in boxes, the pallindromic region is underlined, the coding sequence of the putative gene involved in phage attachment is shown in lower case text, and dots indicate no sequence;

Figure 3 Shows Diagrammatic representation, relative to the chromosome of the cattle type, of the position of SEQ ID NO. 1 and the points of insertion of the tandem repeat and the likely copy of IS*900* in the sheep type; and

Figure 4 PCR products from cattle and sheep types of *M. paratuberculosis* amplified with the three oligonucleotide primers DMC529, DMC531, and DMC533. Lanes: 1 and 11, molecular size markers; 2-5, cattle types; 6-9, sheep types; 10, negative control.

BEST MODES FOR CARRYING OUT THE INVENTION

Experimental

Non-limiting examples illustrating the invention will now be provided. It will be appreciated that the above description is provided by way of example only and variations in the materials and technique used which are known to those skilled in the art are contemplated.

The present invention provides a DNA sequence, SEQ ID NO. 1, that is unique to sheep types of *M. paratuberculosis* and provides the use of SEQ ID NO. 2 for diagnostic testing for organisms of the MAI complex. The present invention also provides for the specific use of SEQ ID NO. 1 and SEQ ID NO. 2 to distinguish between sheep types and cattle types of *M. paratuberculosis* and to distinguish all *M. paratuberculosis* strains from other strains of the closely related MAI complex and from strains of the *M. tuberculosis* complex. A PCR diagnostic test using three oligonucleotide primers is given as an example of the utility of the invention but the invention is not limited to these oligonucleotides or to the use of PCR and a wide range of other diagnostic tests based on these sequences, their complements or any RNA or protein that they specify is envisaged.

Discovery of SEQ ID NO. 1

The strains of the MAI complex used in this work are given in Table 1. Oligonucleotide primers used in the work are given in Table 2. Strains were cultured using standard mycobacterial media (Collins *et al.*, 1997). Purified DNA was extracted as described previously (Collins *et al.*, 1990). When DNA from strains of both the sheep and cattle types of *M. paratuberculosis* was subjected to PCR at an annealing temperature of 50°C using primers DMC136 and DMC137 directed outwards from each end of IS*900*, only DNA from sheep types gave a major product between 300 bp and 400 bp (Figure 1). Subsequently, it was observed that the same 342 bp product was obtained if only one PCR primer (DMC136) was used. The PCR product was extracted from the gel, re-amplified at an annealing temperature of 65°C, cloned into pBluescript KSII (Stratagene) and sequenced (SEQ ID NO. 1). Comparison of this sequence using the programme BLAST against the partially completed sequence of the genome of a cattle type of *M. paratuberculosis* (National Centre for Biotechnology Information database [http://www.ncbi.nlm.nih.gov/]) indicated a high degree of homology to positions 1020 – 1316 of SEQ ID NO. 2, denoted as contig 249 in the database. Comparisons using the GAP programme of GCG (Wisconsin Package Version 10.2, Genetics Computer Group, Madison, Wisconsin) delineated two major differences (Figure 2). First, the sheep type but not the cattle type has a tandem repeat of a 12 bp sequence followed by a 4 bp linker that together contain a 14 bp pallindromic sequence. Second, the cattle type was not homologous to DMC136 at the 5' end. This indicated that an IS*900* element was inserted at the 5' end of SEQ ID NO. 1 in the sheep type but not in the homologous region of SEQ ID NO. 2 in the cattle type. Further investigation using DMC505 and DMC507 indicated that both types have similar sequences at the 3' end because these primers worked equally well in a PCR reaction with an annealing temperature of 60°C on DNA from both sheep and cattle types and gave products differing by only 16 bp. It thus appears that the product shown

in Figure 1 was produced by using the single primer DMC136 because this primer is completely homologous to DNA of the ovine type at the 5' end of SEQ ID NO. 1 (Figure 2) but homologous to only about the last 10 nucleotides at the 3' end. For this reason, the final nine nucleotides of SEQ ID NO. 1 (Figure 2) are shown in lower case text. The most likely explanation for homology of the 5' end of SEQ ID NO. 1 to DMC136 is that a copy of IS900 is inserted at this position in the genome of sheep types but not cattle types and this is shown in diagrammatic form in Figure 3. Confirmation of the presence of this copy of IS*900* was provided by performing a PCR on both sheep and cattle types of *M. paratuberculosis* using the oligonucleotide primers DMC137, which reads out of IS*900* from the opposite end to DMC136, and DMC531 which would be expected to read towards the DMC137 end of IS*900* in sheep types of *M. paratuberculosis*. As expected, this pair of primers gave a product of the expected size with sheep types of *M. paratuberculosis* but no product with cattle types. Clearly, to one with skill in the art this provides alternative regions of IS*900* and SEQ ID NO. 2 to those used in the PCR example below and these alternative regions could be used for designing oligonucleotide primers and constructing PCR tests that potentially have similar utility to that of the example. Further analysis of SEQ ID NO. 1 and SEQ ID NO 2 using other GCG programmes showed that both the tandem repeat and the difference between sheep and cattle types at the 5' end of SEQ ID NO. 1 are in or adjacent to likely coding sequences one of which has high homology to a gene whose product is involved in phage attachment (Barsom and Hatfull, 1996). These differences may therefore be important in determining the host preference of sheep and cattle types. If this is the case, these DNA differences observed between cattle and sheep types may be a very widespread or even ubiquitous phenomenon. Comparison of SEQ ID NO. 1 and SEQ ID NO 2 to the incomplete genome sequence of an *M. avium* subsp. *avium* strain (National Centre for Biotechnology Information database [http://www.ncbi.nlm.nih.gov/]) did not identify any closely homologous sequences apart from the first 285 bp of SEQ ID NO 2. This indicated that SEQ ID NO. 1 and

most of SEQ ID NO. 2 might not be present in *M. avium* subsp. *avium* and that these sequences could thus be used for constructing tests to distinguish between *M. paratuberculosis* and *M. avium* subsp. *avium*.

Development of a PCR assay

A PCR assay was developed using a GeneAmp PCR System 9600 (Applied Biosystems) and the three primers DMC529, DMC531, DMC533 (Table 2 and Figure 2) under the following conditions: 1 cycle at 95°C, 3 min; 25 cycles at 60°C, 30 s, 72°C, 30 s, 94°C, 30 s; 1 cycle at 72°C, 7min. DNA from all 19 strains of the cattle type (Table 1) gave the expected product of 310 bp, while DNA from all 12 strains of the sheep type (Table 1) gave the expected product of 162 bp (Figure 4). A PCR product was not observed for any of the wide range of strains of the MAI complex (Table 1) that did not contain IS*900* and were not *M. paratuberculosis*. No PCR product was observed with strains of *Mycobacterium tuberculosis* and *Mycobacterium bovis* (Table 1) which are members of the *M. tuberculosis* complex. This group of organisms which cause tuberculosis in mammals comprise the following species: *M. tuberculosis*, *M. bovis*, *Mycobacterium bovis* subsp. *caprae*, *Mycobacterium microti* and *Mycobacterium canettii*. *M. tuberculosis* causes most of the tuberculosis in humans and *M. bovis* causes tuberculosis in a wide range of mammals including humans, cattle and deer. In some situations there is diagnostic utility in having a fast diagnostic test such as PCR to distinguish between samples from animals infected with strains of the *M. tuberculosis* complex and those infected with *M. paratuberculosis* (de Lisle *et al.*, 1993).

The MAI complex covers a relatively broad group of genetically related mycobacteria that, with the exception of *M. paratuberculosis*, are found in many environmental niches and are occasional mammalian pathogens. Because of the potential of these organisms to confuse the diagnosis of paratuberculosis, strains of the MAI complex tested in this study were weighted towards those that had been isolated from humans or

from a range of different animal hosts and that might be expected to be most closely related to *M. paratuberculosis* (Collins *et al.*, 1997). The fact that none of these strains was positive in the PCR assay gives a high level of confidence in the utility of the test to distinguish between strains of *M. paratuberculosis* and other strains of the MAI complex. In the case of *M. paratuberculosis*, the inclusion of 10 strains from five other countries, including sheep strains from Canada, South Africa and Iceland, enabled a cross-section of strains with different IS*900* RFLP types to be tested. In all cases, the PCR results were consistent with this RFLP division into sheep and cattle types, indicating that tests based on SEQ ID NO. 1 and SEQ ID NO. 2 should have wide utility for distinguishing between sheep and cattle types of *M. paratuberculosis* in many countries. Because of the association of *M. paratuberculosis* with Crohn's disease in humans, assays based on SEQ ID NO. 1 and SEQ ID NO. 2 not only have the potential to be widely applicable to epidemiological and other studies of paratuberculosis but will also have utility in the field of Crohn's disease.

Table 1. Strains of the MAI complex and *M. tuberculosis* complex subjected to PCR

No. of strains	With IS*900*	With IS*901*	Description	Source
10	0	3	Reference serotypes 1-6, 8-11 of MAI complex*	Dawson **
11	0	7	Cattle, deer and pig isolates of MAI complex*	New Zealand
6	0	5	Bird isolates of MAI complex*	New Zealand
4	0	0	Human isolates of MAI complex*	New Zealand
14	14	–	*M. paratuberculosis* cattle type	New Zealand
3	3	–	*M. paratuberculosis* cattle type	Canada
1	1	–	*M. paratuberculosis* TMC1613; cattle type	USA
1	1	–	*M. paratuberculosis* 316F; cattle type	UK
7	7	–	*M. paratuberculosis* sheep type	New Zealand
1	1	–	*M. paratuberculosis* sheep type	Canada
3	3	–	*M. paratuberculosis* sheep type	South Africa
1	1	–	*M. paratuberculosis* sheep type	Iceland
1	1	–	*M. tuberculosis* H37Rv	USA
2	–	–	*M. bovis*	New Zealand

* None of these strains of the MAI complex were *M. paratuberculosis*
** See Wards *et al.* (1987)

Table 2. DNA oligonucleotide primers used in this work

Oligonucleotide	Sequence 5' – 3'
DMC136	GCTTGACAACGTCATTGAG
DMC137	CCCTTCAAGAAAGGTAAGG
DMC505	CAAGTTGTCGTACTCCTCGTC
DMC507	TTAGCTGACCTATCTACAGGC
DMC529	TTGACAACGTCATTGAGAATCC
DMC531	TCTTATCGGACTTCTTCTGGC
DMC533	CGGATTGACCTGCGTTTCAC

Aspects of the present invention have been described by way of example only and it should be appreciated that modifications and additions may be made thereto without departing from the scope thereof.

AGRESEARCH LIMITED

by their Attorneys

<u>JAMES & WELLS</u>

REFERENCES

Bannantine, J.P., Baechler, E., Zhang, Q., Li, L., Kapur, V. (2002) Genome Scale Comparison of *Mycobacterium avium* subsp. *paratuberculosis* with *Mycobacterium avium* subsp. *avium* reveals potential diagnostic sequences. *J Clin Microbiol*.40: 1303-1310.

Barsom, E. K., and G. F. Hatfull. (1996) Characterization of *Mycobacterium smegmatis* gene that confers resistance to phages L5 and D29 when overexpressed. Mol. Microbiol. 21:159-170.

Beard, P. M., M. J. Daniels, D. Henderson, A. Pirie, K. Rudge, D. Buxton, S. Rhind, A. Greig, M. R. Hutchings, I. McKendrick, K. Stevenson, and J. M. Sharp. (2001) Paratuberculosis infection of nonruminant wildlife in Scotland. J Clin Microbiol 39:1517-1521.

Chiodini, R.J., Van Kruiningen, H.J., Merkal, R.S. (1984) Ruminant paratuberculosis (Johne's disease): the current status and future prospects. *Cornell Vet* 74: 218-262.

Collins, D.M., Gabric, D.M., de Lisle, G.W. (1989) Identification of a repetitive DNA sequence specific to *Mycobacterium paratuberculosis*. *FEMS Microbiol Lett* 60: 175-178.

Collins, D.M., Gabric, D.M., de Lisle, G.W. (1990) Identification of two groups of *Mycobacterium paratuberculosis* strains by restriction endonuclease analysis and DNA hybridization. *J Clin Microbiol* 28: 1591-1596.

Collins, D.M., Stephens, D.M., de Lisle, G.W. (1993) Comparison of polymerase chain reaction tests and faecal culture for detecting *Mycobacterium paratuberculosis* in bovine faeces. *Vet Microbiol* 36: 289-299.

Collins, D.M., Cavaignac, S., de Lisle, G.W. (1997) Use of four DNA insertion sequences to characterise strains of the *Mycobacterium avium* complex isolated from animals. *Mol Cell Probes* 11: 373-380.

de Lisle, G.W., Collins, D.M., Huchzermeyer, H.F.A.K. (1992) Characterization of ovine strains of *Mycobacterium paratuberculosis* by restriction endonuclease analysis and DNA hybridization. *Onderstepoort J Vet Res* 59: 163-165.

de Lisle, G.W., Yates, G.F., Collins, D.M. (1993) Paratuberculosis in farmed deer; case reports and DNA characterization of isolates of *Mycobacterium paratuberculosis*. *J Vet Diagn Invest* 5: 567-571.

Englund, S., Bolske, G., Johansson, K.E.. (2002) An IS*900*-like sequence found in a *Mycobacterium* sp. other than *Mycobacterium avium* subsp. *paratuberculosis*. *FEMS Microbiol Lett* 209: 267-271.

Falkinham, J. O. (1999) Molecular epidemiology,: other mycobacteria. in *Mycobacteria: molecular biology and virulence*, Ratledge, C.R., Dale, J. (eds), Blackwell Science, London, pp. 136-160.

Fang, Y., Wu, W.H., Pepper, J.L,, Larsen, J.L., Marras, S.A., Nelson, E.A., Epperson, W.B., Christopher-Hennings, J. (2002) Comparison of real-time, quantitative PCR with molecular beacons to nested PCR and culture methods for detection of *Mycobacterium avium* subsp. *paratuberculosis* in bovine fecal samples. *J Clin Microbiol*. 40: 287-291.

Grant, I.R., Hitchings, E.I., McCartney, A., Ferguson F., Rowe M.T. (2002) Effect of commercial-scale high-temperature, short-time pasteurization on the viability of *Mycobacterium paratuberculosis* in naturally infected cows' milk. *Appl Envir Microbiol* 68: 602-607.

Green, E.P., Tizard, M.L., Moss, M.T., Thompson, J., Winterbourne, D.J., McFadden, J.J., Hermon-Taylor, J. (1989) Sequence and characteristics of IS*900*, an insertion element identified in a human Crohn's disease isolate of *Mycobacterium paratuberculosis*. *Nucleic Acids Res* 17: 9063-9073.

Harris, N.B., Barletta, R.G. (2001) *Mycobacterium avium* subsp. *paratuberculosis* in veterinary medicine. *Clin Microbiol Rev* 14: 489-512.

Hermon-Taylor, J., Bull, T.J., Sheridan, J.M., Cheng, J., Stellakis, M.L., Sumar, N. (2000) Causation of Crohn's disease by Mycobacterium avium subspecies paratuberculosis. *Can J Gastroenterol* 14: 521-539.

Kunze, Z.M., Wal, S., Appelberg, R., Silva, M.T., Portaels, F., McFadden, J.J. (1991) IS901, a new member of a widespread class of atypical insertion sequences, is associated with pathogenicity in *Mycobacterium avium*. *Mol Microbiol* 5: 2265-2272.

Marsh, I., Whittington, R., Cousins, D. (1999) PCR-restriction endonuclease analysis for identification and strain typing of *Mycobacterium avium* subsp. *paratuberculosis* and *Mycobacterium avium* subsp. *avium* based on polymorphisms in IS*1311*. *Mol Cell Probes* 13: 115-126.

Pavlik, I., Horvathova, A., Dvorska, L., Bartl, J., Svastova, P. du Maine, R., Rychlik, I. (1999) Standardisation of restriction fragment length polymorphism analysis for *Mycobacterium avium* subspecies *paratuberculosis*. *J Microbiol Methods* 38: 155-167.

Thorel, M.F., Krichevsky, M., Levy-Frebault, V.V. (1990) Numerical taxonomy of mycobactin-dependent mycobacteria, emended description of *Mycobacterium avium*, and description of *Mycobacterium avium* subsp. *avium* subsp. nov., *Mycobacterium avium* subsp. *paratuberculosis* subsp. nov., and *Mycobacterium avium* subsp. *silvaticum* subsp. nov. *Int J Syst Bacteriol* 40: 254-260.

Wards, B.J., Collins, D.M., de Lisle, G.W. (1987) Restriction endonuclease analysis of members of the *Mycobacterium avium-M. intracellulare-M. scrofulaceum* serocomplex. *J Clin Microbiol* 25: 2309-2313.

Whittington, R.J., Hope, A.F., Marshall, D.J., Taragel, C.A, Marsh, I. (2000) Molecular epidemiology of *Mycobacterium avium* subsp. *paratuberculosis*: IS*900* restriction fragment length polymorphism and IS*1311* polymorphism analyses of isolates from animals and a human in Australia. *J Clin Microbiol* 38: 3240-3248.

SEQUENCE LISTING

<110> AgResearch Limited
Collins, Desmond Michael

<120> IMPROVEMENTS IN AND RELATING TO A METHOD OF DNA TESTING

<130> 31161/14

<160> 2

<170> PatentIn version 3.1

<210> 1
<211> 342
<212> DNA
<213> Mycobacterium paratuberculosis

<400> 1
gcttgacaac gtcattgaga atccccttgc tacttagctg acctatctac aggctgggtc 60

agcataatcg cagcgtcgga caaggcgtgg aggaatcgcg tcgcgactta tgcaactcct 120

tcaaaaccgg cgtcgcgccg ccgtgtgaaa cgcaggtcaa tccgcaccgg gatggtgatc 180

ggttcaagtc gtccgaccat gggcggttgg ggacccggac actcgagaca tcacccctag 240

ctgatatcag ctttagctga tatcatgcgg tagtgcggaa atgggaggtt gacctgactc 300

tcatcgaagc ctggatggat gccctcaatg acgttgtcaa gc 342

<210> 2
<211> 2306
<212> DNA
<213> Mycobacterium paratuberculosis

<400> 2
ttgaagttgc cgaaggtgcg aatcgagccg tcggtgtcga aggcgacggt gtcgatcttg 60

gtgtgcggca ggcgcaggaa ctcgtcgatg aggcccagct cgtccatgat ccgcagcgtg 120

gtcggatgca cggtgtcgcc gcggaagtcg cgcaggaagt cgttgtgctt ttccaagacg 180

acgaccggga ttccggcccg ggcgagcagc agcgcgtgca ccatcccggc gggaccgccc 240

ccggcaacgc acacctgagt ccggatcact tgcgcaggct acctcccggg tccaaccact 300

aggccgtcca acctcgacgg ccgctacaga gctacacgag ccccgcccga gaacatcgag 360

tcatggaact ctatcgcttc cggcaccgcg cccggcggga tgtcgaacga gacgctccca 420

tccataccga ggccagggtt gatgtcttcc atccaggtgt tgtcgctgag cgaggtcgtc 480

gcgtcaaact tcttgccgtc gatgatcagt ttttggttgg aagcagagaa gctttgggac 540

ttgtcgccga tgttcgtcac gcggagcttc acgacgaaaa actcgccctt ggcttgctgc 600

aaactgaaga ccccttcctt agtggccgag cgatcgacac caagcacttg aaactcgaac 660

ttaccgtctc gcaccgggga cccagcgggg gcgacggcag gccctacatg ctttggcttc 720

gtcggcgcgc ttgcccttga ggggaccgac gctgccgtgt tcgactccgc gacggtcttc 780

ttgtcgccgt cgttgccgcc gctggcgatg gcaatcagga ccaccactac ggctccggcg 840

cctaggatcc acttcttatc ggacttcttc tggcggctgg gcttcacggg ctggccaacg 900

[illegible]gtggcg gcgcgggaaa tggcgcataa ctttcggtcc acgctgtgcc gtcgaagtac 960

cgctggcggc tcggattaga tggatcggga taccaacccg ctgttggctg cgtcatgaag 1020

tccccttgct acttagctga cctatctaca ggctgggtca gcataaccgc agcgtcggac 1080

aaggcgtgga ggaatcgcgt cgcgacttat gcaactcctt caaaaccggc gtcgcgccgc 1140

cgtgtgaaac gcaggtcaat ccgcaccggg atggtgatcg gttcaagtcg tccgaccatg 1200

ggcggttggg gacccggaca ctcgagacat cacccctagc tgatatcatg cggtagtgcg 1260

gaaatgggag gttgacctga ctctcatcga agcctggatg gatgccctcg atgacgagga 1320

gtacgacaac ttgatagcag cgctcgagca gttggaggaa cacggtccaa ttacccggcg 1380

gccgtttgtg gacacccttg aagggtcaag gcacccgaac atgaaggagc tccgtccgcg 1440

ccccacgaaa gctggagccc acattcgcgt gctattcgcc ttcgataccc ggtcgcgggc 1500

gatcatgttg atcgctggag acaaggcggg caattggtcc aagtggtatg ccaagcacat 1560

tcccattgcc gacgagctct tcgatgctca ccaaaagcgt ttacacaagg cagcagccaa 1620

agccaccaat cggaaaccaa ggaagggaaa gaaacgatga ccaatcttga tgacatgcgc 1680

cgtcgccgtc ccggcaaccg gacccgtatc gacgcgatca aggcggagat ggaccgcgag 1740

gttgcccagt accgtctccg cgaactgcgc gaggctgccg gctacacgca gacgactctc 1800

gctgcggcta ttggcgtcgg gcagaaccgc gtctcgcaga tggaacatgg cgatctgggc 1860

acgagccggg tcgacactct ccgcaagtac gtggaagcga ctgggggtga gctggaagtg 1920

tccgtcaagc gtccggacgg gtcacgggtt ctcctcagcc tgtaaagcgc gacgaccgag 1980

acccagggcg aaggcttcga gaacggtgtc caggtaccgg ctgcggcttg tgattctcgg 2040

gtgccctcgg cgctcagaag cagtcgaatg acggggccat cgatacagaa tcgacggaga 2100

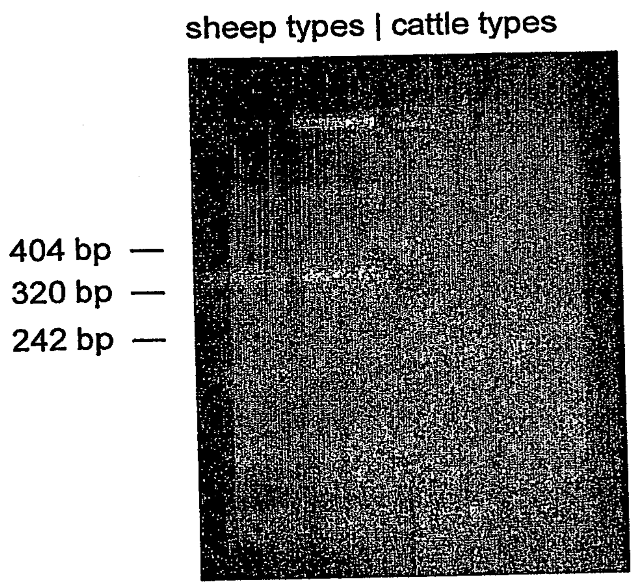
accgaagcag tacaacagat ctcgtcgcac atcggcctca actcgatcgc ggcacagatc 2160

aggtaaccgt attcgcggac gtcgccaaga ccgactccga tgcaacacga tcacccgtct 2220

agcacggcta gagtcacctg tcggactgtc gcgcgaccgc tgtcgtccac cgaagacgac 2280

gaggacacca cggcgaagta ccaccc 2306

FIGURE 1



IGURE 2

```
cattle   850  cacttcttatcggacttcttctggcggctgggcttcacgggctggccaacgccgggtggcggcgcggggaa
   DMC531 ------------------------------------->

cattle   920  atggcgcataactttcggtccacgctgtgccgtcgaagtaccgctggcggctcggattagatggatcggg

                                                    DMC507 ------------------------------------>
cattle   990  ataccaacccgctgttggctgcgtcatGAAGTCCCCTTGCTACTTAGCTGACCTATCTACAGGCTGGGTC
sheep      1  ..........GCTTGACAACGTCATTGAGAATCCCCTTGCTACTTAGCTGACCTATCTACAGGCTGGGTC
                    DMC529 ------------------------------------>
              DMC136 -------------------------------->

cattle  1060  AGCATAACCGCAGCGTCGGACAAGGCGTGGAGGAATCGCGTCGCGACTTATGCAACTCCTTCAAAACCGG
sheep     61  AGCATAATCGCAGCGTCGGACAAGGCGTGGAGGAATCGCGTCGCGACTTATGCAACTCCTTCAAAACCGG

cattle  1130  CGTCGCGCCGCCGTGTGAAACGCAGGTCAATCCGCACCGGGATGGTGATCGGTTCAAGTCGTCCGACCAT
sheep    131  CGTCGCGCCGCCGTGTGAAACGCAGGTCAATCCGCACCGGGATGGTGATCGGTTCAAGTCGTCCGACCAT
                          <-------------------------------- DMC533

cattle  1200  GGGCGGTTGGGGACCCGGACACTCGAGACATCACCCC..................TAGCTGATATCATGCGG
sheep    201  GGGCGGTTGGGGACCCGGACACTCGAGACATCACCCCTAGCTGATATCAGCTTTAGCTGATATCATGCGG

                                                                         <----------------
cattle  1254  TAGTGCGGAAATGGGAGGTTGACCTGACTCTCATCGAAGCCTGGATGGATGCCCTCGATGACGAGGAGTA
sheep    271  TAGTGCGGAAATGGGAGGTTGACCTGACTCTCATCGAAGCCTGGATGGATGCCCTCAATGACGttgtcaa
                                                                   <----------------
              ------------------ DMC505                           3' end of DMC136
cattle  1324  CGACAACTTGATAGC
sheep    341  gc.............
```

FIGURE 3

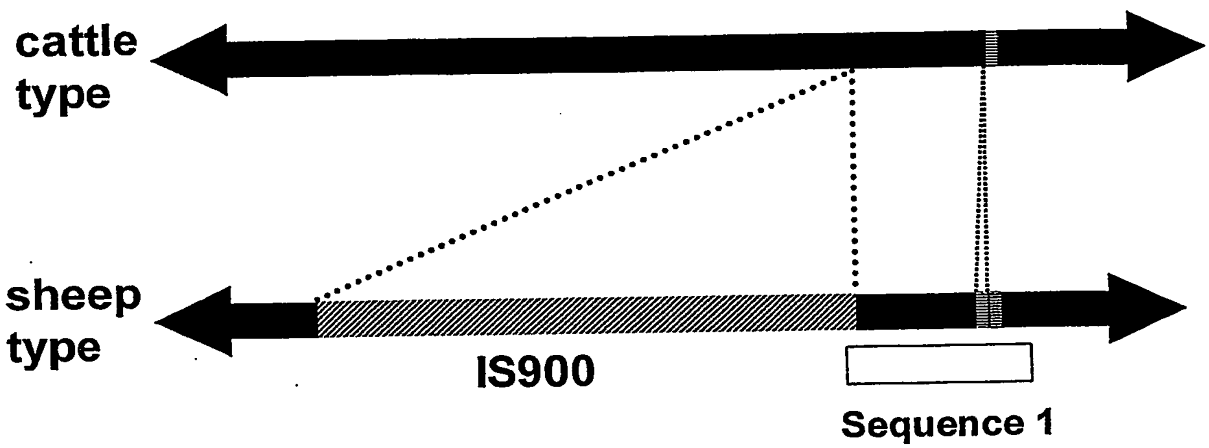


FIGURE 4